

EFFECT OF SENTENTIAL CONTEXT ON SYLLABIC STRESS PERCEPTION BY HEARING-IMPAIRED LISTENERS

Dragana Barac-Cikoja and Sally Revoile

Gallaudet University, Washington, D.C.

ABSTRACT

Although carrier phrases are extensively used in clinical testing of speech recognition, their influence on target identification by hearing-impaired (HI) listeners is poorly understood. The present study examined the effect of sentential context on the identification of the stressed syllable within the target VCV by listeners with mild to severe hearing loss. The study focused on the influence of prosodic structure of the carrier sentence as determined by the location of sentential stress and speaking rate. The target VCV's were presented in contexts that varied in the amount of the prosodic information available about the sentential stress pattern and speaking rate. That is, the target VCV's were presented embedded either in a sentence, in a noise simulation of the sentence, or extracted from the sentence and presented in isolation. Results indicated that both the prosodic context and the presentation condition affected performance. When sentential stress was on the target VCV, performance accuracy was highest. The finding is linked to the acoustic salience of the stressed syllable. The effect of the presentation condition is discussed in terms of the hearing-impaired listeners' possible difficulties in segmenting the sentence and allocating attention.

INTRODUCTION

Perception of lexical stress has been studied minimally for hearing-impaired persons. For persons with normal hearing, the redundant acoustic cues of syllable stress enable facile identification of stressed versus unstressed syllables. However, hearing-impaired persons may have greater difficulty in perceiving lexical stress. Due to the auditory reception deficiencies of these persons, the acoustic cues of syllable stress might be less salient than for normal-hearing persons. Especially for hearing-impaired persons with more extensive hearing loss, the acoustic characteristics differentiating stressed versus unstressed syllables could be indiscriminable. Such reduced prosodic information may also limit the efficiency of hearing-impaired listeners in attending to the segments that carry greater information load within a speech stream. A related effect could be a reduced ability to perceive the consonant in a VCV. Recent research (Revoile et al., 1995) showed that hearing-impaired listeners identified the consonant in VCV's better when the VCV's were presented in isolation than when the acoustically identical VCV's were embedded in a carrier sentence. The authors suggested that this performance difference may be due to the hearing-impaired listeners' difficulties in parsing the words of continuous speech.

In the present paper, we report on a part of an ongoing study that investigates the effect of sentential context on the perception of syllable stress by hearing-impaired listeners. The accuracy of the syllabic stress identifications by hearing-impaired listeners was examined under conditions that differed: first, in the prosodic structure of the sentence to which the perceptual target belonged; and second, in the amount of the prosodic information present in the target's context. The strength of a syllabic stress can be expected to vary depending on context. For example, the sentence type, the presence and relative position of the emphatic stress in the sentence, or speech rate, can all influence acoustic salience of the syllabic stress. In the stimulus corpus that was used in this investigation, either location of the sentential (emphatic) stress or the speaking rate were manipulated, and the effects on the accuracy of the syllabic stress location within the target VCV were analyzed. Listeners' performances were compared under three presentation methods: the target VCV embedded in the sentence - providing unaltered information on the sentential stress pattern; the target VCV embedded in a noise carrier that preserved only temporal and amplitude characteristics of the original sentence; and, the target VCV extracted from the sentence and presented in isolation without any speech-like context. We asked: (1) Does position of the sentential stress relative to the target VCV affect syllabic stress identification? (2) Does identification of stressed syllable differ for fast versus slow utterances? (3) If significant, does the effect of each of these two factors depend on the prosodic information provided by the target's context?

METHOD

Subjects. Nine hearing-impaired adults from the Gallaudet University community participated as paid listeners. As defined by their pure tone threshold averages (PTA, 0.5, 1.0, and 2.0 kHz), the listeners' hearing losses were mild ($n = 1$, PTA 30 dB HL), moderate ($n = 4$, PTA 41 to 65 dB HL); or severe ($n = 4$, PTA 70 to 85 dB HL).

Stimuli. The core set of stimuli were 80 utterances of "You put /VCV/ to bed", among which the /VCV/ was produced with stress on V1 ($n = 40$) or on V2 ($n = 40$). Either of two variables was manipulated in subsets (40 sentences each) of the core stimuli: (1) sentence speaking rate, or (2) location of the emphatic stress in the sentence; i.e., either YOU, or PUT, or VCV, or BED was emphasized. For the latter condition, the sentence was spoken at the speaker's comfortable speaking rate (4.4 syl/s). When speaking rate was the independent variable, all 40 utterances had sentential stress on the VCV. The sentences were produced at either a slow rate (2.5 syl/s), or a faster rate (4.5 syl/s).

These core stimuli composed the baseline condition, from which two other conditions were generated: (1) Extracted VCV's, and (2) Noise simulation of the sentence; i.e., gaussian noise low-pass filtered at 1.0 kHz, replaced the entire sentence except for the VCV. The noise was amplitude shaped to preserve the amplitude envelope and timing characteristics of the words in the sentence.

Procedure. Listeners identified whether V1 or V2 was stressed for each stimulus presented in a single-interval trial. No feedback was provided to the listeners. The three presentation conditions were administered in separate blocks. Five administrations of a stimulus block were tested, each on different days. The listeners choose a most comfortable level at which to hear the VCV's. A one-minute break was given at the half-way point of each block.

RESULTS

For each of the 18 conditions (two speaking rates plus four sentential stress conditions x three presentation methods [VCV in sentence, in noise carrier, or extracted]), percent information transmitted (%IT) was computed on the basis of listeners correct and incorrect syllable stress identification. The group of listeners' %IT scores were submitted to further analyses. Separate repeated measures analyses of variance (ANOVA) for each of the sentential stress and speech rate manipulations were conducted. In the sentential stress ANOVA, significant main effects of both sentential stress location, $F(3, 24) = 9.63$ ($p < .001$), and presentation method, $F(2, 16) = 5.10$ ($p < .05$) were found. Pairwise comparisons (Helmert contrasts) revealed that the significant sentential stress effect was due to the improved performance when the VCV received the primary stress in the sentences. The group's %IT scores when the VCV was stressed were better by about 20% than for stress on any of the other words. No differences among the other three sentential stress locations reached significance. The average %IT for the four sentential stress locations with presentation method as a parameter are presented in Figure 1.

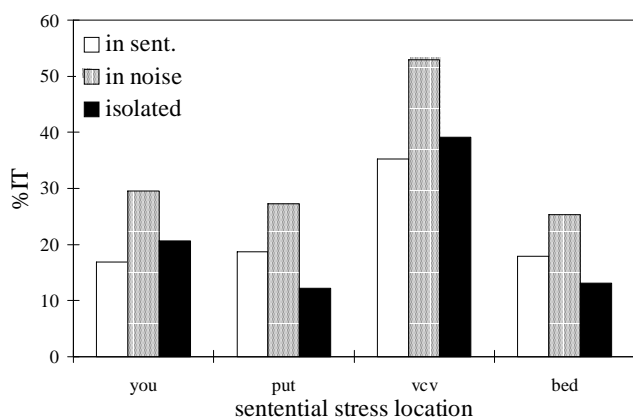


Figure 1. Group mean %IT as a function of the sentential stress location and the presentation method.

The significant effect of presentation method was due to easier stress identification when the VCV was presented in the noise carrier than when in either sentential context or extracted and presented in isolation. The interaction between sentential stress location and presentation method was not found significant ($F[6, 48] = 1.04, p > .1$). However, inspection of the Figure 1 suggests, and the pairwise comparisons confirmed, that in the extracted VCV condition, performance was better when the sentence of origin had emphasis on YOU than when the primary stress was on the words contiguous to the VCV, that is, on PUT and BED. This trend was not observed when VCV's were embedded in either sentential or noise context.

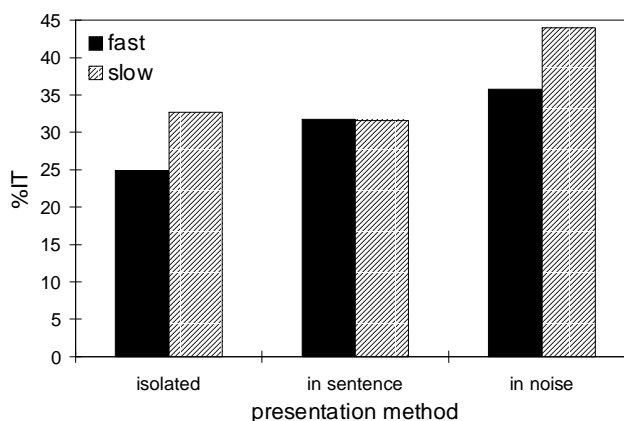


Figure 2. Group mean %IT as a function of speaking rate and presentation method.

The ANOVA for the speech rate manipulations showed a significant effect of presentation method, $F(2, 16) = 5.59$ ($p < .05$). However, neither the main effect of rate nor the rate by presentation method interaction reached significance ($p > .1$). Figure 2 shows the average %IT for the three presentation condition with speech rate as a parameter. The advantage of the noise carrier over the other presentation methods is again, apparent.

DISCUSSION

The first questions asked in this study can be answered positively: Identification of the stressed syllable by hearing-impaired people was affected by the manipulation of the sentential stress position. The stress identification was most accurate when the primary sentential stress was on the target VCV. This was true for each of the presentation methods, that is, for the target presented in isolation, embedded in the sentence, or embedded in the carrier noise. The accuracy of the listeners' performance on the other three sentential stress conditions (emphasis either on YOU, PUT or BED), did not vary significantly for this group of subjects.

The effect of the emphatic stress position can be accounted for by the greater acoustic salience of the stressed syllable when the sentential emphasis was on the target VCV than when it was on

any other word in the sentence. Results of the analyses of duration and amplitude (RMS dB) of the stressed relative to the unstressed vowel in the target VCV's (measures often used to express temporal and amplitude prominence, respectively, of the stressed syllable) will illustrate this claim. Relative duration of the stressed syllable was an average of 21% larger when sentential stress was on the VCV as compared to the average VCV duration when the other words in the sentence were stressed. For relative amplitude, a 90% average difference was found between the stressed syllables in the VCV's with versus without sentential stress. Our finding on the effect of sentential prosody on the acoustic salience of syllabic stress is consistent with the related literature (e.g., Waibel, 1986).

The speaking rate manipulations as represented in our stimulus set, were not found significant in any of the presentation conditions. Although the temporal difference between the fast and slow productions was substantial, it is possible that the clear-speech constraint that was imposed on the speaker, and the requirement for the full vowel /a/ in both positions within VCV in the stimulus selection process, biased the effect of rate on the stress characteristics.

Finally, this study revealed that across speaking rate and sentential stress manipulations, hearing-impaired listeners identified syllable stress in VCV's better when VCV's were in a noise carrier (simulating a sentence) than when in the sentence of origin, or when presented extracted from the sentence. Presenting the target in the sentence as opposed to presenting it extracted from the sentence was not detrimental to the hearing-impaired listeners' performances. Although opposite to the findings of the consonant identification task (Revoile et al., 1995), lack of advantage for the extracted VCV's is not surprising given the task at hand. Syllabic stress is a context dependent value. While the extracted VCV provided some context to determine stress, it may have not been sufficient. From the whole sentence, and from the neighboring syllables in particular, a listener could use information to establish the relative stress value of the syllables in the VCV. Further, it has been hypothesized that in English, a stress-timed language, rhythmic pattern established by stressed syllables may facilitate efficient use of limited attentional capacity by allowing anticipatory allocation of attention (Pitt & Samuel, 1990). Thus, if the presence of the sentential content imposed segmentation difficulty, it was clearly counteracted by the useful prosodic information from the context.

Consistent with the attention hypothesis is the finding that the presentation in the amplitude-shaped noise led to the best performance. While it eliminated (or at least greatly reduced) the segmentation difficulty, the noise carrier contained temporal and amplitude information about the sentential stress pattern useful in determining the relative value of the syllable stress.

For hard of hearing persons, an implication from this effect is relevant for clinical tests of word recognition. Such tests typically involve presentation of a target word at or near the end of a carrier phrase, e.g., "You will say___". If a hard of hearing

person under test has difficulty segmenting or parsing the carrier phrase, a consequence could be interference in locating the target word. Based on this study, we would suggest replacing the carrier phrase with noise that simulates the temporal and amplitude characteristics of the phrase. The resultant prosodic syllabic patterns in the noise may facilitate focus on the target word by the hard of hearing person, whose auditory attentional capacity may be constrained by a hearing deficit.

REFERENCES

1. Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 564-573.
2. Revoile, S., Kozma-Spytek, L., Holden-Pitt, L., Pickett, J., & Droge, J. (1995). Acoustic-phonetic context considerations for speech recognition testing of hearing-impaired listeners. *Ear & Hearing*, **16**, 254-262.
3. Waibel, A. (1986). Recognition of lexical stress in continuous speech understanding system - A pattern recognition approach. *International Conference on Acoustics, Speech and Signal Processing*, 2287-2290.